

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: HUSE, WILLIAM D.

(ii) TITLE OF INVENTION: SURFACE EXPRESSION LIBRARIES OF HETEROMERIC RECEPTORS

(iii) NUMBER OF SEQUENCES: 75

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: PRETTY, SCHROEDER, BRUEGEMANN & CLARK
 (B) STREET: 444 SO. FLOWER STREET, SUITE 200
 (C) CITY: LOS ANGELES
 (D) STATE: CALIFORNIA
 (E) COUNTRY: UNITED STATES
 (F) ZIP: 90071

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
 (B) FILING DATE:
 (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: CAMPBELL, CATHRYN A.
 (B) REGISTRATION NUMBER: 31,815
 (C) REFERENCE/DOCKET NUMBER: P31 8882

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 619-535-9001
 (B) TELEFAX: 619-535-8949

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7445 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: both
 (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTCAAG CTCGGCCCC AAATGAAAAT 60
 ATAGCTAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120
 CGTTCCAGA ATTGGGAATC AACTGTTACA TCGGAATGAAA CTTCCAGACCA CGGTACTTTA 180
 GTTGCATATT TAAAACATGT TGAGCTACAG CACCAAGATTC AGCAATTAAG CTGTAAGCCA 240
 TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAGG TACTCTCTAA TCCTGACCTG 300
 TTGGAGTTG CTTCCGGTCT GGTTGCGTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG 360
 TCTTTGGGC TTCCCTCTAA TCTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420

2nd
C3

CAGGGTAAAG ACCTGATTIT TGATTATGG TCATTCTCGT TTTCTGAAC TTTAAAGCA 480
 TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540
 AAACATTTA CTATTACCCC CTCTGGAAA ACTTCTTTG CAAAAGCCTC TCGCTATTTT 600
 GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCCTGGT 660
 AATTCTTTT GGC GTTATGATCTGCATTA GTTGAATGTG GTATTCCCTAA ATCTCAACTG 720
 ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT 780
 TCTTCCCAAC GTCCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAAATTCA 840
 CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGT 900
 CTCGTCAAGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG 960
 AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC 1020
 TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC 1080
 GTCTGGCCT CGTTCCGGCT AAGTAACATC GAGCAGGTGG CCGATTTCGA CACAATTAT 1140
 CAGGGGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT 1200
 CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCGCTTTCTGT TTTAGGTTGG TGCCTTCGTA 1260
 GTGGCATTAC GTATTTTACCGT CGTTAAATGG AAACCTCCCTC ATGAAAAAGT CTTTAGTCCT 1320
 CAAAGCCTCT GTAGCCGTTG CTACCCCTCGT TCCGATGCTG TCTTCGCTG CTGAGGGTGA 1380
 CGATCCCGCA AAAGCGGCCT TTAACCTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA 1440
 TGCCTGGCGC ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA 1500
 ATTCAACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCCTTT GGAGCCTTTT 1560
 TTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTGCAA TTCCCTTAGT TGTTCCCTTC 1620
 TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTAGCAA AACCCCATAC AGAAAATTCA 1680
 TTTACTAACG TCTGGAAAGA CGACAAAATC TTAGATCGTT ACGCTAACTA TGAGGGTTGT 1740
 CTGTGGAATG CTACAGGGCT TGAGTTTGT ACTGGTGACG AAACCTCAGTG TTACGGTACA 1800
 TGGGTTCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT 1860
 TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATAACACCT 1920
 ATTCCGGGCT ATACTTATAT CAACCCCTCTC GACGGCACTT ATCCGCTGG TACTGAGCAA 1980
 AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT 2040
 CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACGT TTTATACGGG CACTGTTACT 2100
 CAAGGCAGTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG 2160
 TATGACGGCTT ACTGGAACCG TAAATTGAGA GACTGGCTT TCCATTCTGG GTTTAATGAA 2220
 GATCCATTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCGTGTCAAT 2280
 GCTGGCGGGCG GCTCTGGTGG TGGTTCTGGT GCGGGCTCTG AGGGTGGTGG CTCTGAGGGT 2340
 GGC GGTTCTG AGGGTGGCGG CTCTGAGGGA GGC GGTTCCG GTGGTGGCTC TGGTCCGGT 2400
 GATTTGATT ATGAAAAGAT GGCAAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT 2460

CAAAACGGCG TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCC GTGACGGTGA TAATTACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT	2700
TTTGTCTTTA CGCCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760
TTCCCGTGGTG TCCTTGCGTT TCCTTTATAT GTGCCACCT TTATGTATGT ATTTTCTAAG	2820
TTTGCTAACAA TACTGCCGTAATTAAGGAGCT TAATCATGCC AGTTCTTTG GGTATTCCGT	2880
TATTATTGCG TTTCCCTCGGT TTCCCTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTCATT GTTGTGTTGCT CTTATTATTG	3000
GGCTTAACTC AATTCTTGTG GGTTATCCT CTGATATTAG CGCTCAATTAA CCCTCTGACT	3060
TTGTTCAAGGG TGTTCAAGTTA ATTCTCCCGT CTAATGCCCT TCCCTGTTT TATGTTATTG	3120
TCTCTGTAAA GGCTGCTATT TTCACTTTG ACGTTAAACA AAAAATCGTT TCTTATTG	3180
ATTGGGATAAA ATAATATGGC TGTTTATTG GIAAGTGGCA AATTAGGCTC TGGAAAGAGC	3240
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT	3360
CTTACAATAC CGGATAAGCC TTCTATATCT GATTGCTTG CTATTGGCG CGGTAATGAT	3420
TCCTACGATG AAAATAAAA CGGCTTGCCTT GTTCTCGATG AGTGGGGTAC TTGGTTAAT	3480
ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTCT ACATGCTCGT	3540
AAATTAGGAT GGGATATTAT TTTCTTGTGTT CAGGACTTAT CTATTGTTGA TAAACAGGGC	3600
CGTTCTGCAT TAGCTGAACA TGTTGTTAT TGTCGTCGTC TGGACAGAAAT TACTTTACCT	3660
TTTCTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT	3720
GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT	3780
ACTCGTAAGA ATTGTATAA CGCATATGAT ACTAACACAGG CTTTTCTAG TAATTATGAT	3840
TCCGGTGTGTT ATTCTTATTAA AACGCCCTAT TTATCACACG GTCGGTATTT CAAACGATTA	3900
AATTAGGTC AGAAGATGAA GCTTACTAAA ATATATTGA AAAAGTTTC ACGCGTTCTT	3960
TGTCCTGGCA TTGGATTGCA ATCAGCATT ACATATAGTT ATATAACCCAA ACCTAAGCCG	4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTGATA AATTCACTAT TGACTCTCT	4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAAATTAT	4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTATG TACTGTTCC	4200
ATTAAGGAG GTAAATTCAA TGAAATTGTT AAATGTAAATT AATTGTTGTT TCTTGATGTT	4260
TGTTTCTATCA TCTTCTTTG CTCAGGTAAT TGAAATGAAT AATTCCCTC TGCGCGATT	4320
TGTAACCTGG TATTCAAAGC AATCAGGCCA ATCCGTTATT GTTTCTCCCG ATGTAAGG	4380
TACTGTTACT GTATATTCACT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTATTTC	4440
TGTTTACGT GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAAGAAGTA	4500

TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA	4560
TGATAATTCC GCTCCCTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC	4620
TTTAAATT AATAACGTTG GGGCAAAGGA TTTAATACGA GTTGTGAAAT TGTTTGTAAA	4680
GTCTAACTACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT	4740
TAGTGCACCT AAAGATATT TAGATAACCT TCCTCAATTG CTTTCTACTG TTGATTTGCC	4800
AACTGACCAG ATATTGATTG AGGGTTGAT ATTTGAGGTG CAGCAAGGTG ATGCTTTAGA	4860
TTTTCATTT CCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTGCTCGGT ATTTTAATG GCGATGTTT	4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTAT	5100
TACTGGTCGT GTGACTGGTG AATCTGCAA TGTAAATAAT CCATTTAGA CGATTGAGCG	5160
TCAAAATGTA GGTATTCCTA TGAGCGTTT TCCCTGTTGCA ATGGCTGGCG GTAATATTGT	5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA AGAAGTATTG CTACAACGCT TAATTTGGGT GATGGACAGA CTCTTTACT	5340
CGGTGGCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCGTCTAA	5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTAA AGCCGGCGG	5520
GTGTGGTGGT TACGGCGAGC GTGACCGCTA CACTTGGCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGGCTTCTT CCCTTCCTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTAGGGTTC CGATTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG	5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTT CGCCCTTTGA	5760
CGTTGGAGTC CACGGTCTTT AATAGTGGAC TCTTGTCTCA AACTGGAACA ACACTCAACC	5820
CTATCTCGGG CTATTCTTT GATTATAAG GGATTGGCC GATTTGGAA CCACCATCAA	5880
ACAGGATTT CGCCTGCTGG GGCAAACCGAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGGCGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCC	6000
GGCCGCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCAATTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGAA CGCAATTAAAT GTGAGTTAGC	6120
TCACTCAATTAA GGCACCCAG GCTTACACT TTATGCTTCC GGCTCGTATG TTGTTGGAA	6180
TTGTTGAGGGATAACAAATTTC CACACGGCTC ACTTGGCACT GGCCGTCGTT TTACAACGTC	6240
GTGACTGGGA AAACCCCTGGC GTTACCCAAG CTTTGTACAT GGAGAAAATA AAGTGAAACA	6300
AAGCACTATT GCACTGGCAC TCTTACCGTT ACCGTTACTG TTTACCCCTG TGACAAAAGC	6360
CGCCCAAGTC CAGCTGCTCG AGTCAGGCCT ATTGTGCCA GGGGATTGTA CTAGTGGATC	6420
CTAGGCTGAA GGCGATGACC CTGCTAAGGC TGCATTCAAT AGTTTACAGG CAAGTGGTAC	6480
TGACTACATT GGCTACGCTT GGGCTATGGT AGTACTTATA GTTGGTGTCA CCATAGGGAT	6540

TAAATTATTC AAAAGTTA CGAGCAAGGC TTCTTAAGCA ATAGCCAAGA GGGCCGGCACC	6600
GATCGCCCTT CCCAACAGTT CGCGAGCCTG AATGGCGAAT GGGCCTTGC CTGGTTTCCC	6660
GCACCAAG CGGTCCCGGA AAGCTGGCTG GAGTGGATC TTCCCTGAGGC CGATACGGTC	6720
GTCGTCCTT CAAACTGGCA GATGCACGGT TACGATGCC CGATCTACAC CAACGTAACC	6780
TATCCCATT A CGGTCAATCC GCGTTTGTT CCCACGGAGA ATCCGACGGG TTGTTACTCG	6840
CTCACATT A ATGTTGATGA AAGCTGGCTA CAGGAAGGCC AGACGGAAT TATTTTGAT	6900
GGCGTTCTA TTGGTTAAAA AATGAGCTGA TTTAACAAA ATTTAACGCG AATTTAAACA	6960
AAATATTAAC GTTTACAATT TAAATATTG CTTATACAAT CTTCCCTGTT TTGGGGCTTT	7020
TCTGATTATC AACCGGGGTA CATATGATTG ACATGCTAGT TTTAQQGATTA CCGTTCATCG	7080
ATTCTCTTGT TTGCTCCAGA CTCTCAGGCA ATGACCTGAT AGCCTTGTA GATCTCTCAA	7140
AAATAGCTAC CCTCTCCGGC ATTAATTTAT CAGCTAGAAC CGTTGAATAT CATATTGATG	7200
GTGATTGAC TGTCTCCGGC CTTCTCACC CTTTGAAATC TTTACCTACA CATTACTCAG	7260
GCATTGCATT TAAAATATAT GAGGGTTCTA AAAATTTTA TCCTTGCCTT GAAATAAAGG	7320
CTTCTCCGGC AAAAGTATT A CAGGGTCATA ATGTTTTGG TACAACCGAT TTAGCTTTAT	7380
GCTCTGAGGC TTTATTGCTT AATTTGCTA ATGTTTGCC TTGCTGTAT GATTTATTGG	7440
ACGTT	7445

(2) INFORMATION FOR SEQ ID NO:2:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7317 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTCA GCTGGCCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTGGAGA ATTGGGAATC AACTGTTACA TGGAAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCA GATTGAGT AGCAATTAAAG CTCTAAGGCC	240
TCCGAAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTTGGCTTT GAAGCTCGAA TTAAAACGGG ATATTTGAAG	360
TCTTTCGGGC TTCTCTCTAA TCTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTG TGATTATGG TCATTCTCGT TTCTGAAC TTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTATGAC GATTCCGAG TATTGGACGC TATCCAGTCT	540
AAACATTTA CTATTACCCC CTCTGGAAA ACTTCTTTG CAAAAGCCTC TCGCTATTG	600
GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGGCTCGT	660
AATTCCCTTT GCGGTTATGT ATCTGCATTA GTTGAATGTG GTATTCTAA ATCTCAACTG	720

ATGAATCTT	CTACCTGAA	TAATGTTGTT	CCGTTAGTTC	GTTTTATTAA	CGTAGATTTT	780
TCTTCCCAAC	GTCCTGACTG	GTATAATGAG	CCAGTTCTTA	AAATCGCATA	AGGTAATTCA	840
CAATGATTAA	AGTTGAAATT	AAACCATCTC	AAGCCCAATT	TACTACTCGT	TCTGGTGT	900
CTCGTCAGGG	CAAGCCTTAT	TCACTGAATG	AGCAGCTTG	TTACGTTGAT	TTGGGTAAATG	960
AATATCCGGT	TCTTGTCAAG	ATTACTCTTG	ATGAAGGTCA	GCCAGCCTAT	GGCCCTGGTC	1020
TGTACACCGT	TCATCTGTCC	TCTTCAAAG	TTGGTCAGTT	CGGTTCCCTT	ATGATTGAEC	1080
GTCTGCGCCT	CGTTCCGGCT	AAAGAACATG	GAGCAGGTGCG	CGGATTTCGA	CACAATTAT	1140
CAGGGCAGTA	TACAAATCTC	CGTTGTACTT	TGTTTCCGGC	TTGGTATAAT	CGCTGGGGGT	1200
CAAAGATGAG	TGTTTTAGTG	TATTCTTCG	CCTCTTCGT	TTAGGTTGG	TGCCCTCGTA	1260
GTGGCATTAC	GTATTTTAC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTAGTCCT	1320
CAAAGCCTCT	GTAGCCGTTG	CTACCCCTCGT	TCCGATGCTG	TCTTCGCTG	CTGAGGGTGA	1380
CGATCCCGCA	AAAGCGGCCT	TTAACCTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTIA	1440
TGCGTGGCGG	ATGGTTGTTG	TCATTGTCGG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA	1500
ATTACCTCG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTGGAGA	TTTCAACGT	GAACAAATT	TTATTCCAA	TTCTCTTAGT	TGTTCCCTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTAGCAA	AAACCCATAC	AGAAAATTCA	1680
TTTACTAACG	TCTGGAAAGA	CGACAAAAT	TTAGATCGTT	ACGCTAACTA	TCAGGGTTGT	1740
CTGTGGAATG	CTACAGGGT	TCTAGTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCCCTCTC	GACGGCACTT	ATCCGCTGG	TACTGACCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCCGAAA	TAGGCAGGGG	GCATTAACCT	TTTATACGGG	CACTGTTACT	2100
CAAGGCACTG	ACCCCGTAA	AACTTATTAC	CACTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTGAGA	CACTGGCTT	TCCATTCTGG	TTTAATGAA	2220
GATCCATTCTG	TTTGTGAATA	TCAAGGCCAA	TCTGCTGACC	TGGCTCAACC	TCCTGTCAAT	2280
GCTGGCGGCG	GCTCTGGTGG	TGGTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCCGTTCTG	AGGGTGGCGG	CTCTGAGGGG	GGCGGTTCCG	GTGGTGGCTC	TGGTCCGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	ATAAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
AAAAACGCGC	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTGCGTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTCC	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTACCT	TCCCTCCCTC	AATCGGTTGA	ATGTGCCCT	2700
TTTGTCTTAA	GGCGCTGGTAA	ACCATATGAA	TTTCTATTG	ATTGTGACAA	AATAAACTTA	2760

TTCCGTGGTG	TCTTGCCTT	TCTTTATAT	GTTGCCACCT	TTATGTATGT	ATTTCTACG	2820
TTTCTAACA	TAATGCCGAA	TAAGGAGTCT	TAATCATGCC	AGTTCTTGTG	GGTATTCCGT	2880
TATTATTGCG	TTTCCTCGGT	TTCCCTCTGG	TAACCTTGTG	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CTATTCATT	GTTTCTTGCT	CTTATTATTG	3000
GGCTTAACTC	AATTCCTGTC	GGTTATCTCT	CTGATATTAG	CGCTCAATT	CCCTCTGACT	3060
TTGTTCAAGG	TGTTCAAGTTA	ATTCTCCCGT	CTAATGCCGT	TCCCTGTTT	TATGTTATTC	3120
TCTCTGTAAA	GGCTGCTATT	TTCATTTTG	ACGTTAAACA	AAAAATCGTT	TCTTATTG	3180
ATTGGGATAA	ATAATATGGC	TGTTTATTTT	GTAACTGGCA	AATTAGGCTC	TGGAAAGACG	3240
CTCGTTAGCG	TTGGTAAGAT	TCAGGATAAA	ATTGTAGCTG	GGTGCAAAAT	AGCAACTAAT	3300
CTTGATTTAA	GGCTTCAAAAA	CCTCCCGCAA	GTCGGGAGGT	TCGCTAAAC	GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC	TTCTATATCT	GATTTGCTTG	CTATTGGCG	CGGTAATGAT	3420
TCCTACGATG	AAAATAAAAAA	CGGCTTCCTT	GTTCTCGATG	AGTGGCGTAC	TTGGTTAAAT	3480
ACCCGTTCTT	GGAATGATAA	GGAAAGACAC	GGGATTATTG	ATTGGTTTCT	ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT	TTTCTTGTT	CAGGACTTAT	CTATTGTTGA	AAAACAGGCG	3600
CGTTCTGCAT	TAGCTGAACA	TGTTGTTTAT	TGTCGTCGTC	TGGACAGAAT	TACTTTACCT	3660
TTTGTGGTA	CTTTATATTTC	TCTTATTACT	GGCTCGAAAAA	TGCCCTGCG	AAAATTACAT	3720
CTTGGCGTTG	TTAAATATGG	CGATTCTCAA	TTAAGCCCTA	CTGTTGAGCG	TTGGCTTTAT	3780
ACTGGTAAGA	ATTGTATAA	CCGATATGAT	ACTAAACAGG	CTTTTCTAG	TAATTATGAT	3840
TCCGGTGT	TTTCTTATTT	AACGCCCTAT	TTATCACACG	GTCGGTATTT	CAAACCATTA	3900
AATTAGGTC	AGAAGATGAA	GCTTACTAAA	ATATATTGA	AAAAGTTTC	ACCGGTTCTT	3960
TGTCTTGCAT	TTGGATTTC	ATCAGGATT	ACATATAGTT	ATATAACCCA	ACCTAAGCCG	4020
GAGGTTAAAA	AGGTAGTCTC	TCAGACCTAT	GATTTGATA	AATTCACTAT	TGACTCTTCT	4080
CAGCGTCTTA	ATCTAACGCTA	TCGCTATGTT	TTCAAGGATT	CTAAGGAAA	ATTAATTAAAT	4140
ACCGACGATT	TACAGAACCA	AGGTTATTCA	CTCACATATA	TTGATTATG	TACTGTTCC	4200
ATTAAGGAAAG	GTAAATTCAA	TGAAATTGTT	AAATGTAATT	AATTGTTT	TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTG	CTCAGGTAAT	TGAAATGAAT	AATTGCCCTC	TGCGCGATT	4320
TGTAACCTGG	TATTCAAAGC	AATCAGGCCA	ATCCGTTATT	GTTTCTCCCG	ATGAAAAGG	4380
TACTGTTACT	GTATATTCA	CTGACGTTAA	ACCTGAAAAT	CTACGCAATT	TCTTATTTC	4440
TGTTTTACGT	GCTAATAATT	TTGATATGGT	TGGTTCAATT	CCCTCCATAA	TTCAGAAAGT	4500
TAATCCAAAC	AATCAGGATT	ATATTGATGA	ATTGCCATCA	TCTGATAATC	AGGAATATGA	4560
TGATAATTCC	GCTCCTTCTG	GTGGTTCTT	TGTTCCGCAA	AATGATAATG	TTACTCAAAC	4620
TTTAAAAATT	AAATAACGTT	GGGCAAAGGA	TTAATACGA	GTGTCGAAT	TGTTTGAA	4680
GTCTAATACT	TCTAAATCCT	CAAATGTATT	ATCTATTGAC	GGCTCTAATC	TATTAGTTGT	4740
TAGTGCACCT	AAAGATATTT	TAGATAACCT	TCCTCAATT	CTTCTACTG	TTGATTGCC	4800

AACTGACCAAG ATATTGATTG AGGGTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTT	4860
TTTTCAATT GCTGCTGGT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTATCTT CTGCTGGTGG TTGCTTCGGT ATTTTAATG GCGATGTTT	4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATTCTTACCG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTAT	5100
TACTGGTGGT GTGACTGGTG AATCTGCCAA TGTAAATAAT CCATTTCAGA CGATTGAGCG	5160
TCAAAATGTA GGTATTTCCA TGAGCGTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT	5220
TCTGGATAATT ACCAGCAAGG CCGATAGTTT GAGTTCTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA AGAAGTATTG CTACACCGT TAATTGCGT GATGGACAGA CTCTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA	5400
AATCCCTTA ATCGGCCTCC TGTTAGGTC CGCGCTGTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGGTGTGTC GTCAAAGCAA CCATAGTAGG CGCCCTGTAG CGGGCGCATTA AGCGCGGCGG	5520
GTGTGGTGGT TACGGCGAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGCTTTCTT CCCTTCCCTT CTGGCCACGT TCGGGGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTAGGGTTC CGATTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG	5700
ATTTGGGTGA TGGTTACCGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTGTG	5760
CGTTGGAGTC CACGTTCTT AATAGTGGAC TCTTGTCCA AACTGGAACA ACACCTCAACC	5820
CTATCTCGGG CTATTCTTT GATTATAAG GGATTITGCC GATTTCGGAA CCACCATCAA	5880
ACAGGAIITTT CGCCTGCTGG GCCAAACCAAG CGTGGACCCC TTGCTGCAAC TCTCTCAGGG	5940
CGAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCC	6000
GGCGCCCAAT ACGAAACCG CCTCTCCCCG CGCGTTGCC GATTCAATTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAAT GTGAGTTAGC	6120
TCACTCATTAA GGCAACCCAG CTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTT CACACGCCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC	6240
TACGGCAGCC GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCG AGCTCGTGT	6300
GACCCAGACT CCAGATATCC AACAGGAATG AGTGTAAATT CTAGAACCGG TCACTTGGCA	6360
CTGGCCGTGG TTTTACAACG TCGTGACTGG GAAAACCTG CGCTTACCCA AGCTTAATCG	6420
CTTTCAGAA TTCCCTTCTG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATGCC	6480
TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCTT GCCTGGTTTC CGGCACCAGA	6540
AGCGGTGCCG GAAAGCTGGC TGGAGTGCAGA TCTTCCTGAG GCGGATACGG TCGTCGTCCC	6600
CTCAAACCTGG CAGATGCACG GTTACGATGC GCCCATCTAC ACCAACGTAA CCTATCCC	6660
TACGGTCAAT CCCCCCTTTC TTCCCAACCGA GAATCCGACG GCTTGTACT CGCTCACATT	6720
TAATGTTGAT GAAAGCTGGC TACAGGAAGG CCAGACGCCA ATTATTTTG ATGGCGTTCC	6780
TATTGGTTAA AAAATGAGCT GATTAAACAA AAATTTAACG CGAATTTAA CAAAATATTA	6840

ACGTTTACAA TTAAATATT TGCTTATACA ATCTTCCTGT	TTTGGGCT TTTCTGATTA	6900
TCAACCGGGG TACATATGAT TGACATGCTA	GTTCACGAT TACCGTCAT CGATTCTCTT	6960
GTTCCTCCA GACTCTCAGG CAATGACCTG	ATAGCCTTG TAGATCTCTC	7020
ACCCCTCTCCG GCATTAATT	ATCAGCTAGA ACGGTTGAAT	7080
ACTGCTCTCGG GCCTTTCTGA	CCCTTTGAA TCTTTACCTA	7140
TTTAAATAT ATGAGGGTTC	CACATTACTC AGGCATTGCA	7200
GCAAAAGTAT TACAGGGTCA	TAATGTTTT GGTACAACCG	7260
GCTTATTGCG TTAATTGTC	ATTTAGCTTT ATGCTCTGAG	7317
TAATTCTTGT	CCTTGCCTGT ATGATTATT GGATGTT	

(2) INFORMATION FOR SEQ ID NO:3:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7729 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATGCTACTA CTATTAGTAG AATTGATGCC	ACCTTTTCAG CTCGGCCCC AAATGAAAAT	60
ATAGCTAAC AGGTATTGCA	CCATTGCGA AATGTATCTA ATGGTCAAAC	120
CGTTCGCAGA ATTGGGAATC	AACTGTTACA TCGGAATGAAA	180
GTGCAATT TAAAACATGT	TGAGCTACAG CACCAAGATT	240
TCTGCAAAAA TGACCTCTTA	CAATTAAAGG TACTCTCTAA	300
TTGGAGTTTG CTICCGGTCT	GGTCGCTTT GAAGCTCGAA	360
TCTTCGGGC TTCTCTTAA	TTAAAACGCG ATATTGAAAG	420
CAGGGTAAAG ACCTGATTTT	TGATTATGG TCATTCTCGT	480
TTTGAGGGGG ATTCAATGAA	TTTCTGAACG TGTAAAGCA	540
AAACATTTTA CTATTACCCC	TTTCTGAG TATTGAC	600
GGTTTTATC GTCGCTGGT	CAAAAGCCTC TCGCTATT	660
AATTCTTTT GGCGTTATGT	AAACGAGGGT TATGATAGTG	720
ATGAATCTTT CTACCTGTAA	TTGCTCTTAC TATGCCTCGT	780
TCTTCCCAAC GTCTGACTG	AAATCCGATA AGGTAAATTCA	840
CAATGATTAA AGTTGAAATT	AAACCATCTC AAGCCCAATT	900
CTCGTCAGGG CAAGCCTTAT	TACTACTCGT TCTGGTGT	960
AATATCCGGT TCTTGTCAAG	TTACGTTGAT TTGGTAAATG	1020
TGTACACCGT TCATCTGTCC	TTGGTCAGTT CGGTTCCCTT	1080
GTCTGCGCCT CGTTCCGGCT	ATGAAAGGTCA	1140
AGTAACATG GAGCAGGTCG	GGGATTTCGA CACAATTAT	
CAGGGATGA TACAAATCTC	TGTTGTACTT TGTTTGC	1200
	CCCTGGGGGT	

CAAAGATGAG	TCTTTAGTG	TATTCTTCG	CCTCTTCGT	TTTAGGTTGG	TGCCTTCGTA	1260
GTGGCATTAC	GTATTTTAC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTTAGTCCT	1320
CAAAGCCTCT	GTAGCCGTTG	CTACCCCTCGT	TCCGATGCTG	TCTTCGCTG	CTGAGGGTGA	1380
CGATCCCGCA	AAAGCCGCT	TTAACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTTA	1440
TGCGTGGCG	ATGGTTGTTG	TCATTGTCGG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA	1500
ATTCACCTCG	AAAGCAAGCT	GATAAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTT	1560
TTTTGGAGA	TTTTCACCGT	AAAAAAATTA	TTATTCGCAA	TTCCCTTAGT	TGTTCCCTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTAGCAA	AACCCCATAAC	AGAAAATTCA	1680
TTTACTAACG	TCTGGAAAGA	GGACAAAAACT	TTAGATCGTT	ACGGTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGGT	TGTAGTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCCCTCTC	GACGGCACIT	ATCCGCCTGG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGTAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	GGTTCCGAAA	TAGGCAGGGG	GCATTAACGT	TTTATACGGG	CACTGTTACT	2100
CAAGGCAC TG	ACCCCGTTAA	AACTTATTAC	CA GTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACGCTT	ACTGGAACGG	TAAATTCA	GACTGCCCTT	TCCATTCTGG	CTTTAATGAA	2220
GATCCATTG	TTTGTGAATA	TCAAGGCCAA	TCGTCTGACC	TGCCCTCAACC	TCCTGTCAAT	2280
GCTGGCGGG	GCTCTGGTGG	TGGTTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTTCCGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	ATAAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
GAAAACGCC	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCGCTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCCTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTGG	GTGACGGTGA	TAATTCACCT	2640
TTAATGAATA	ATTTCCGTCA	ATATTTACCT	TCCCTCCCTC	AATCGGTGA	ATGTCGCCCT	2700
TTTGTCTTTA	GCGCTGGTAA	ACCATATGAA	TTTCTATTG	ATTGTGACAA	AATAAAACTTA	2760
TTCCGTGGTG	TCTTTGCGTT	TCTTTTATAT	GTGCCCCCT	TTATGTATGT	ATTTTCTACG	2820
TTTGCTAAC	TACTGGTAA	TAAGGAGTCT	TAATCATGCC	AGTTCTTTG	GGTATTCCGT	2880
TATTATTGCG	TTTCCCTCGGT	TTCCCTCTGG	TAACTTTGTG	GGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAAG	ATAGCTATTG	CTATTTCATT	GTTCCTTGCT	CTTATTATTG	3000
GGCTTAAC	AATTCTTGTG	GGTTATCTCT	CTGATATTAG	CGCTCAATT	CCCTCTGACT	3060
TTGTTCA	GGTTCA	ATTCTCCCGT	CTAATGGCCT	TCCCTGTTT	TATGTTATT	3120
TCTCTGTAAA	GGCTGCTATT	TTCA	ACGTTAAACA	AAAAATGTT	TCTTATTG	3180
ATTGGGATAA	ATAATATGGC	TGTTTATT	GTAACTGGCA	AATTAGGCTC	TGGAAAGACG	3240

CTCGTTAGCG	CTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTTAA	GCCTCAAAA CCTCCGCAA CTCGGGAGGT TCGCTAAAAC GCCTCGCGTT	3360
CTTAGAATAC	CGGATAAGCC TTCTATATCT GATTTGCTT CTATTGGCG CGGTAATGAT	3420
TCCTACGATG	AAAATAAAAA CGGCTGCTT GTTCTCGATG AGTGGGTAC TTGGTTAAT	3480
ACCCGTTCTT	GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTCT ACATGCTCGT	3540
AAATTAGGAT	GGGATATTAT TTTTCTTGT CAGGACTTAT CTATTGTTCA TAAACAGGCG	3600
CGTTCTGCAT	TAGCTGAAGA TGTGTTTAT TGTCTCGTC TGGACAGAAT TACITTTACCT	3660
TTTGTGGTA	CTTTATATTG TCTTATTACT GGCTCGAAA TGCCTCTGCC TAAATTACAT	3720
GTTGGCGTTG	TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCC TTGGCTTAT	3780
ACTGGTAAGA	ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTCTAG TAATTATGAT	3840
TCCGGTGT	TTTCTTATTG AAAGCCTTAT TTATCACACG GTGGTATTT CAAACCATT	3900
AATTAGGTC	AGAAGATCAA GCTTACTAAA ATATATTGA AAAAGTTTC ACGCGTTCTT	3960
TGTCTTGC	TTGGATTTCG ATCAGGATT ACATATAGTT ATATAACCCA ACCTAACGCCG	4020
GAGGTTAAA	AGGTAGTCTC TCAGACCTAT GATTTGATA AATTCACTAT TGACTCTTCT	4080
CAGCGTCTT	ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTATG TACTGTTTCC	4200
ATTA	AAAAAG GTAAATTCAA TGAAATTGTT AAATGTAATT AATTTGTTT TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTG CTCAGGTAAT TGAAATGAAT AATTCCCTC TGCGCGATT	4320
TGTAAC	TTGG TATTCAAAGC AATCAGGCCA ATCGTTATT GTTCTCCCG ATGTAAGG	4380
TACTGTTACT	GTATATTCACT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTATTTC	4440
TGTTTACGT	GCTAATAATT TTGATATGGT TGGTCAATT CCTCCATAA TTCAAGGTA	4500
TAATCCAAAC	AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA	4560
TGATAATTCC	GCTCCTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC	4620
TTTAA	AAATAACGTTG GGGCAAAGGA TTTAATACGA GTGTCGAAT TGTTTGTAAA	4680
GTCTAATACT	TCTAAATCCT CAAATGTATT ATCTATTGAC GGGCTAAATC TATTAGTTGT	4740
TAGTGCACCT	AAAGATAATT TAGATAACCT TCCTCAATTG CTTTCTACTG TTGATTGCG	4800
AACTGACCA	AGATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTACA	4860
TTTTGATTT	GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GCGGGTGTAA ATACTGACCG	4920
CCTCACCTCT	GTTTATCTT CTGCTGGTGG TTGTTCCGGT ATTTTAATG CGCATGTTT	4980
AGGGCTATCA	GTTCGGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATTCTTACG	CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTAT	5100
TACTGGTCGT	GTGACTGGTG AATCTGCCAA TGTAAATAAT CCATTCAGA CGATTGAGCG	5160
TCAAAATGTA	GGTATTCCA TGAGCGTTT TCCTGTTGCA ATGGCTGGCG GTAAATATTGT	5220
TCTGGATATT	ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCCA GTGATGTTAT	5280

TAATAATCAA AGAAAGTATTG CTACAACGGT TAATTTGGGT GATGGACAGA CTCTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA	5400
AATCCCTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATT ACGCGGGCGG	5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGCTTCTT CCCTTCCTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTTAGGGTTC CGATTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG	5700
ATTCGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTT CGCCCTTGA	5760
CGTTGGAGTC CACGTTCTT AATAGTGGAC TCTTGTCCA AACTGGAACA ACACTCAACC	5820
CTATCTCGGG CTATTCTTT GATTATAAG GGATTGGCC GATTTCGGAA CCACCATCAA	5880
ACAGGATTT CGCCTGCTGG GGCAAACCAAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTAAAAGAA AAACCACCC	6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCACTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCGACTGG AAAGGGGGCA GTGAGGCCAA CGCAATTAAAT GTGAGTTAGC	6120
TCACTCACTT GGCACCCCAAG GCTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGAA	6180
TTGTGAGCGG ATAACAATTT CACACGGCTC ACTTGGCACT GGCGTCTGTT TTACAACGTC	6240
GTGACTGGGA AAACCTGGC GTTACCCAAG CTTGTACAT GGAGAAAATA AAGTGAACAA	6300
AAGCACTATT GCACCTGGCAC TCTTACCGTT ACTGTTTACC CCTGTGGCAA AAGCCCAGGT	6360
CCAGCTGCTC GAGTCGGTCT TCCCCCTGGC ACCCTCTCC AAGAGCACCT CTGGGGGCAC	6420
AGCGGCCCTG GGCTGCCTGG TCAAGACTAA TTCCCCGAAC CGGTGACGGT GTCGTGGAAC	6480
TCAGGGCCCC TGACCAGGGG CGTGCACACC TTCCCCGGCTG TCCTACAGTC CTCAGGACTC	6540
TACTCCCTCA GCAGCGTGGT GACCGTGGCC TCCAGCAGCT TGGGCACCCA GACCTACATC	6600
TCCAACGTCA ATCACAAGCC CAGCAACACC AAGGTGGACA AGAAAGCAGA GCCCAAATCT	6660
TGTACTAGTG GATCCTACCC GTACCGACGTT CCGGACTACGG CTTCTTAGGC TGAAGGGCAT	6720
GACCCCTGCTA AGGCTGCATT CAATAGTTA CAGGCAAGTG CTACTGAGTA CATTGGCTAC	6780
GCTTGGCTA TGGTAGTAGT TATAGTTGGT GCTACCATAG GGATTAATT ATTCAAAAG	6840
TTTACGAGCA AGGCTCTTA AGCAATAGCG AAGAGGGCCCG CACCGATCGC CCTTCCCAAC	6900
AGTTGGCAG CCTGAATGGC GAATGGCGT TTGGCTGGTT TCCGGCACCA GAAGGGTGC	6960
CGGAAAGCTG GCTGGAGTGC GATCTTCTG AGGGCGATAAC GGTGCTCGTC CCCTCAAAC	7020
GGCAGATGCA CGGTTACGGAT CGGCCCCATCT ACACCAACGTT AACCTATCCC ATTACGGTCA	7080
ATCCGGCGTT TGTTCCCACG GAGAATCCGA CGGGTTGTTA CTCGCTCACA TTAAATGTTG	7140
ATGAAAGCTG GCTACAGGAA GGCGAGACGG GAATTATTT TGATGGCGTT CCTATTGGTT	7200
AAAAAAATGAG CTGATTTAAC AAAAATTAA CGCGAATTAA AACAAAATAT TAACGTTTAC	7260
AATTAAATA TTGCTTATA CAATCTTCTG TTTTTGGGG CTTTCTGAT TATCAACCGG	7320

GGTACATATG ATTGACATGC TAGTTTACG ATTACCGTTC ATCGATTCTC TTGTTTGCTC	7380
CAGACTCTCA GGCATGACC TGATAGCCCTT TGTAGATCTC TCAAAAATAG CTACCCCTCTC	7440
CGGCATTAAT TTATCAGCTA GAACGGTTGA ATATCATATT GATGGTGATT TGACTGTCTC	7500
CGGCCTTCT CACCCATTG AATCTTAC TACACATTAC TCAGGCATTG CATTAAAAT	7560
ATATGAGGGT TCTAAAAATT TTTATCCTTG CGTTGAAATA AAGGCTTCTC CGGCAAAAGT	7620
ATTACAGGGT CATAATGTT TTGGTACAAC CGATTTAGCT TTATGCTCTG AGGCTTTATT	7680
GCTTAATTGTT GCTAATTCTT TGCCTTGCT GTATGATTG TTGGACGTT	7729

(2) INFORMATION FOR SEQ ID NO:4:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7557 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATGCTACTA CTATTAGTAG AATTGATGOC ACCTTTCAAG CTCGGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTTGGCA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAAATGAAA CTTCCAGACA CCCTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAAGATT AGCAATTAAG CTCTAAGCCA	240
TCCGAAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG	360
TCTTCGGGC TTCCCTTAA TCTTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTIT TGATTTATGG TCATTCTCGT TTTCTGAAC GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGGAG TATTGGACGC TATCCAGTCT	540
AAACATTTA CTATTACCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCCTGGT	660
AATTCTTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGAA TAATGTTGTT CCGTTAGTTG GTTTTATTAA CGTAGATTIT	780
TCTTCCCAAC GTCCGTACTG GTATAATGAG CCAGTTCTTA AAATGCCATA AGGTAATTCA	840
CAATGATTAA AGTGAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTIT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTGT TTACGTCAT TTGGTAATG	960
AATATCCGGT TCTTGCAAG ATTACTCTTG ATGAAGGTCA GCCAGCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC	1080
GTCTGGCCCT CGTTCCGGCT AAGTAACATG GAGCAGGTGCG CGGATTTCCA CACAATTAT	1140
CAGGCCATGA TACAAATCTC CGTTGACTT TGTTTGGCC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTGTG TATTCTTCTG CCTCTTTCTG TTTAGGTTGG TGCCTTCGTA	1260

GTGGCATTAC	GTATTTIACC	CGTTTAATGG	AAACTTCCTC	ATGAAAAAGT	CTTAGTCCT	1320
CAAAGCCTCT	GTACCCGTG	CTACCCCTCGT	TCCGATGCTG	TCTTTCGCTG	CTGAGGGTGA	1380
CGATCCCGCA	AAAGCGGCCT	TTAACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTTA	1440
TGGCTGGGCG	ATGGTTGTG	TCATTGTCGG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA	1500
ATTCACCTCG	AAAGCAAGCT	GATAAACCGA	TACAATTAAA	GGCTCCTTTT	GGAGCCTTTT	1560
TTTTGGAGA	TTTCAACGT	GAAAAAATTAA	TTATTGGAA	TTCCCTTAGT	TGTTCCCTTC	1620
TATTCTCACT	CCGCTGAAAC	TGTTGAAAGT	TGTTTAGCAA	AACCCCATAC	AGAAAATTCA	1680
TTTACTAACG	TCTGGAAAGA	CGACAAAACT	TTAGATCGTT	ACGCTAACTA	TGAGGGTTGT	1740
CTGTGGAATG	CTACAGGGT	TGTAGTTTGT	ACTGGTGACG	AAACTCAGTG	TTACGGTACA	1800
TGGGTTCTA	TTGGGCTTGC	TATCCCTGAA	AATGAGGGTG	GTGGCTCTGA	GGGTGGCGGT	1860
TCTGAGGGTG	GCGGTTCTGA	GGGTGGCGGT	ACTAAACCTC	CTGAGTACGG	TGATACACCT	1920
ATTCCGGGCT	ATACTTATAT	CAACCCCTCTC	GACGGCACTT	ATCCGCTGG	TACTGAGCAA	1980
AACCCCGCTA	ATCCTAATCC	TTCTCTTGAG	GAGTCTCAGC	CTCTTAATAC	TTTCATGTTT	2040
CAGAATAATA	CGTTCCGAAA	TAGGCAGGGG	GCATTAACGT	TTTATACGGG	CACTGTTACT	2100
CAAGGCACTG	ACCCCGTTAA	AACTTAAATAC	CACTACACTC	CTGTATCATC	AAAAGCCATG	2160
TATGACCGCTT	ACTGGAACGG	TAAATTCAAGA	CACTGCGCTT	TCCATTCTGG	CTTTAATGAA	2220
GATCCATTCTG	TTTGTGAATA	TCAAGGCCAA	TGGTCTGACC	TGCCTCAACC	TCCTGTCAAT	2280
GCTGGCGGGCG	GCTCTGGTGG	TGGTCTGGT	GGCGGCTCTG	AGGGTGGTGG	CTCTGAGGGT	2340
GGCGGTTCTG	AGGGTGGCGG	CTCTGAGGGA	GGCGGTTCCG	GTGGTGGCTC	TGGTCCGGT	2400
GATTTTGATT	ATGAAAAGAT	GGCAAACGCT	AATAAGGGGG	CTATGACCGA	AAATGCCGAT	2460
GAAAACGCCG	TACAGTCTGA	CGCTAAAGGC	AAACTTGATT	CTGTCGCTAC	TGATTACGGT	2520
GCTGCTATCG	ATGGTTTCAT	TGGTGACGTT	TCCGGCCTTG	CTAATGGTAA	TGGTGCTACT	2580
GGTGATTTTG	CTGGCTCTAA	TTCCCAAATG	GCTCAAGTGG	GTGACGGTGA	TAATTACACT	2640
TTAATGAATA	ATTTCCGTCA	ATATTTACCT	TCCCTCCCTC	AATGGTTGA	ATGTCGCCCT	2700
TTTGTCTTTA	GGCGTGGTAA	ACCATATGAA	TTTCTATG	ATTGTGACAA	AATAAACTTA	2760
TTCCGTGGTG	TCTTTCGGTT	TCTTTATAT	GTGCCCACCT	TTATGTATGT	ATTTCTAGG	2820
TTTGCTAACG	TACTGCGTAA	TAAGGAGTCT	TAATCATGCC	AGTTCTTTG	GGTATTCCGT	2880
TATTATTGCG	TTTCCCTCGGT	TTCCCTCTGG	TAACCTTGTT	CGGCTATCTG	CTTACTTTTC	2940
TTAAAAAGGG	CTTCGGTAAG	ATAGCTATTG	CCTGTTCTT	GCTCTTATTAA	TTGGGCTTAA	3000
CTCAATTCTT	GTGGGTTATC	TCTCTGATAT	TAGCGCTCAA	TTACCCCTG	ACTTTGTICA	3060
GGGTGTTCAAG	TTAATTCTCC	CGTCTAATGCC	GCTTCCCTGT	TTTATGTAA	TTCTCTCTGT	3120
AAAGGCTGCT	ATTTTCATTT	TTGACGTTAA	ACAAAAAAATC	TTTCTTATTAA	TGGATTGGGA	3180
TAAATAATAT	GGCTGTTTAT	TTTGTAACTG	GCAAATTAGG	CTCTGGAAAG	ACGCTCGTTA	3240
GGCTGGTAA	GATTCAAGGAT	AAAATTGTAG	CTGGGTCCAA	AATAGCAACT	AACTTGTATT	3300

TAAGGCTTCA	AAACCTCCCC	CAAGTCGGGA	GGTTCGCTAA	AACGCCCTCGC	GTTCTTAGAA	3360
TACCGGATAA	GCCTTCTATA	TCTGATTTCG	TTGCTATTGG	GCGCGGTAAT	GATTCCCTACG	3420
ATGAAAATAA	AAACGGGTTG	CTTGTCTCG	ATGAGTCCGG	TACTTGGTTT	AATACCCGTT	3480
CTTGGAAATGA	TAAGGAAAGA	CAGCCGATT	TTGATTGGTT	TCTACATGCT	CGTAAATTAG	3540
GATGGGATAT	TATTTTCTT	GTTCAAGGACT	TATCTATTGT	TGATAAACAG	GCGCGTTCTG	3600
CATTAGCTGA	ACATGTGTT	TATTGTCGTC	GTCTGGACAG	AATTACTTTA	CCTTTGTCG	3660
GTACTTATA	TTCTCTTATT	ACTGGCTCGA	AAATGCCCT	GCCTAAATT	CATGTTGGCG	3720
TTGTTAAATA	TGGCGATTCT	CAATTAAGCC	CTACTGTTGA	GCCTTGGCTT	TATACTGGTA	3780
AGAATTGTA	TAACGCATAT	GATACTAAAC	AGGCTTTTC	TAGTAATTAT	GATTCCGGTG	3840
TTTATTCTTA	TTTAACGCC	TATTTATCAC	ACGGTCGGTA	TTTCAAACCA	TTAAATTAG	3900
GTCAGAAGAT	GAAGCTTA	ACTAAATT	TGAAAAAGTT	TTCACGGCTT	CTTGTCTTG	3960
CGATTGGATT	TGCATCAGCA	TTTACATATA	GTATATAAC	CCAACCTAAG	CCGGAGGTTA	4020
AAAAGGTAGT	CTCTCAGACC	TATGATTTG	ATAAATTACAC	TATTGACTCT	TCTCAGGGTC	4080
TTAATCTAAG	CTATCGCTAT	GTTTCAAGG	ATTCTAAGGG	AAAATTAAATT	AAATGCCGACG	4140
ATTTACAGAA	GCAAGGTAT	TCACTCACAT	ATATTGATT	ATGTA	TGTTTCA	4200
AAGGTAAATT	AAATGAAATT	GTAAATGTA	ATAAATT	TTTCTTGT	GTGTTTCA	4260
TCATCTTCTT	TTGCTCAGGT	AATTGAAATG	AATAATTGCG	CTCTGCGCGA	TTTTGTA	4320
TGGTATTCAA	AGCAATCAGG	CGAATCCGTT	ATTGTTCTC	CCGATGTAAA	AGGTACTGTT	4380
ACTGTATATT	CATCTGACGT	AAAACCTGAA	AATCTACGCA	ATTTCTTAT	TTCTGTTT	4440
CGTGCTAATA	ATTTGATAT	GGTGGTTCA	ATTCTTCCA	TAATTCA	GTATAATCCA	4500
AACAATCAGG	ATTATATGCA	TGAATTGCCA	TCATCTGATA	ATCAGGAATA	TGATGATAAT	4560
TCCGCTCCTT	CTGGTGGTTT	CTTGTCTCCG	CAAATGATA	ATGTTACTCA	AACTTTAAA	4620
ATTAATAACG	TTCGGGAAA	GGATTTAATA	CGAGTTGTCG	AAATTGTTGT	AAAGTCTAAT	4680
ACTTCTAAAT	CCTCAAATGT	ATTATCTATT	GACGGCTCTA	ATCTATTAGT	TGTTAGTGCA	4740
CCTAAAGATA	TTTGTGATAA	CCTTCGTCAA	TTCTTCTGTA	CTGTTGATTT	GGCAACTGAC	4800
CAGATATTGA	TTGAGGGTTT	GATATTGAG	TTTCAGCAAG	GTGATGCTTT	AGATTTTCA	4860
TTTGCTGCTG	GCTCTCAGCG	TGGCACTGTT	GCAGGCGGTG	TTAATACTGA	CCGGCTCACC	4920
TCTGTTTAT	CTTCTGCTGG	TGGTTCGTT	GGTATTTTA	ATGGCGATGT	TTTAGGGCTA	4980
TCAGTTCCGG	CATTAAGAC	TAATAGCCAT	TCAAAATAT	TGTCTGTC	ACGTATTCTT	5040
ACGGTTTCAG	GTCAGAAGGG	TTCTATCTCT	GTTGGCCAGA	ATGCCCCTT	TATTACTGGT	5100
CGTGTGACTG	GTGAATCTGC	CAATGTAAAT	AATCCATTTC	AGACGATTGA	GGGTCAAAT	5160
GTAGGTATT	CCATGAGCGT	TTTCTCTGTT	GCAATGCC	GGGTAATAT	TGTTCTGGAT	5220
ATTACCAGCA	AGGCCGATAG	TTTGAGTTCT	TCTACTCAGG	CAAGTGTGATGT	TATTACTAAT	5280
CAAAGAACTA	TTGCTACAAC	GGTTAATTG	CGTGATGGAC	AGACTCTTT	ACTCGGTGGC	5340

CTCACTGATT ATAAAAACAC TTCTCAAGAT TCTGGCGTAC CGTTCTGTGTC TAAAATCCCT 5400
 TTAATCGGCC TCCTGTAG CTCCCGCTCT GATTCCAACG AGGAAAGCAC GTTATACGTG 5460
 CTCGTCAAAG CAACCATACT ACAGCGCCCTG TAGGGGGCA TTAAGGGCGG CGGGTGTGGT 5520
 GGTTACGCCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA CGGCCCGCTC CTTTCGCTTT 5580
 CTTCCCTTCC TTCTCGCCA CGTTGCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT 5640
 CCCTTAGGG TTCCGATTAA CTGCTTACG GCACCTCGAC CCCAAAAAAC TTGATTGEGG 5700
 TGATGGTCA CGTAGTGGGC CATGCCCTG ATAGACGGTT TTTCGCCCTT TGACGTTGGA 5760
 GTCCACGGTTC TTAAATAGTG GACTCTTGTG CCAAACCTGGA ACAACACTCA ACCCTATCTC 5820
 GGGCTATTCT TTTCATTTAT AAGGGATTTT GCGGATTTCG GAACQACCAT CAAACAGGAT 5880
 TTTCCCTGCG TGGGGCAAAAC CAGCGTGGAC CGCTTGCTGC AACTCTCTCA GGGCCAGGCG 5940
 GTGAAGGGCA ATCAGCTGTT GCGCGTCTCG CTGGTAAAAA GAAAAACAC CCTGGCGCCC 6000
 AATACGCAAA CGGCCCTCC CGCGCGTTC GCGGATTCTAT TAATGCAGCT GGCACGACAG 6060
 GTTTCCCGAC TGGAAAGCGG GCAGTGAGCC CAACGCAATT AATGTGAGTT AGCTCACTCA 6120
 TTAGGCACCC CAGGCTTAC ACTTTATGCT TCGCTCGT ATGTTGTGTG GAATTGTGAG 6180
 CGGATAACAA TTTCACACGC CAAGGAGACA CTGATAATGA AATACCTATT GCCTACGGCA 6240
 GCGCGTGGAT TGTATTACT CGCTGCCAA CCAGCCATGG CCGAGCTCTT CCCGCCATCT 6300
 GATGAGCAGT TCAAATCTGG AACTGCCCTCT GTTGTGTGCC TGCTGAATAA TTCTATCCC 6360
 AGAGAGGCCA AAGTACAGTG GAAGGTGGAT AACGCCCTCC AATCGGGTAA CTCCCAGGAG 6420
 AGTGTACAG ACCAGGACAG CAAGGACAGC ACCTACAGCC TCAGCAGCAC CCTGACGCTG 6480
 AGCAAAGCAG ACTACGGAGAA ACACAAAGTC TACGCCCTCG AAGTCACCCA TCAGGGCTG 6540
 AGCTCGCCCG TCACAAAGAG CTTCAACAGG GGAGAGTGTG CTAGAACCGC TCACTTGGCA 6600
 CTGGCCGTGCG TTTTACAACG TCGTGACTGG GAAAACCTG GCGTTACCCA AGCTTAATCG 6660
 CCTTGCAGAA TTCCCTTTCG CCAGCTGGCG TAATAGCCAA GAGGCCCGCA CCGATGCC 6720
 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGGTTT GCGTGGTTTC CGGCACCAAGA 6780
 AGCGGTGCCG GAAAGCTGGC TGGAGTGGCA TCTTCCTGAG GCGGATAACGG TCGTCGTCCC 6840
 CTCAAACCTGG CAGATGCCAGG GTTACCGATGC GCCCATCTAC ACCAACGTAA CCTATCCCAT 6900
 TACGGTCAAT CGGCCGTTC TTCCCAACGGA GAATCCGACG GGTTGTTACT CGCTCACATT 6960
 TAATGTTGAT GAAAGCTGGC TACAGGAAGG CCAGACCGGA ATTATTTTG ATGGGGTTT 7020
 TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTAA CAAAATATTA 7080
 ACCTTACAA TTAAATATT TGCTTATACA ATCTTCCTGT TTTTGGGGCT TTCTGATTA 7140
 TCAACCGGGG TACATATGAT TCACATGCTA GTTTACGAT TACCGTCTCAT CGATTCTCTT 7200
 GTTTGCTCCA GACTCTCAGG CAATGACCTG ATAGCCTTTC TAGATCTCTC AAAAATAGCT 7260
 ACCCTCTCCG GCATTAATTG ATCAGCTAGA ACGGTTGAAT ATCATATTGA TCGTGATTTG 7320
 ACTGTCTCCG GCCTTCTCA CCCTTTGAA TCTTACCTA CACATTACTC AGGGATTGCA 7380

TTTAAAATAT ATGAGGGTC TAAAAATTT TATCCTTGGC TTGAAATAAA GGCTTCTCCC	7440
GCAAAAGTAT TACAGGGTCA TAATGTTTT GGTACAACCG ATTTAGCTTT ATGCTCTGAG	7500
GCTTTATTGC TTAATTTTGC TAATTCTTTC CCTTGCCTGT ATGATTATT GGATGTT	7557

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8118 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATCCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGGCCCC AAATGAAAAT	60
ATAGCTAAC AGGTTATTGA CCATTGSGA ATGCTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCCGAGA ATTGGGAATC AACTGTTACA TCGAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAAGATTC AGCAATTAAG CTCTAAGCCA	240
TCTGAAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCC TGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTTCGCTT GAAGCTCGAA TTAAAACCGG ATATTTGAAG	360
TCTTCCGGGC TTGCTCTTAA TCTTTTGAT GCAATCCGCT TTGCTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTIT TGATTTATGG TCATTCTCGT TTTCTGA ACT GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGGAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGAAA ACTTCTTTG CAAAAGCCTC TCGCTATTIT	600
GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCTCGT	660
AATTCCCTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTG GTTTTATTAA CGTAGATTIT	780
TCTTCCCAAC GTCCCTGACTG GTATAATGAG CCAGTCTTA AAATCCGATA AGGTAATTCA	840
CAATGATTAA AGTTGAAATT AAACGATCTC AAGCCCATT TACTACTCGT TCTGGTGTIT	900
CTCGTCAGGG CAAGCCTTAT TCACGTAAAG AGCAGCTTGT TTACGTTGAT TTGGTAAATG	960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTCAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC	1080
GTCTGGCCCT CGTTCCGGCT AAGTAACATG GAGCAGGTGG CGGATTTCGA CACAATTAT	1140
CAGGGCATGA TACAAATCTC CGTTGTACTT TGTTTCCGGC TTGGTATAAT CGCTGGGGT	1200
CAAAGATGAG TGTTTGTG TATTCTTCCG CCTCTTCCGT TTTAGGTTGG TGCCTTGTAA	1260
GTGGCATTAC GTATTTTAC CGTTTAATGG AAACCTCCTC ATGAAAAAGT GTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTG CTACCCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCCGGCT TTAACCTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440
TGGCTGGGGC ATGGTTGTG TGATCTGTGG CGCAACTATC GGTATCAAGC TGTTTAAAGAA	1500

ATTACACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTT GGAGCCTTT 1560
 TTTTGGAGA TTTTCAACGT GAAAAAATTAA TTATTCGCAA TTCCCTTAGT TGTTCCCTTC 1620
 TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTAGCAA AACCCCATAAC AGAAAATTCA 1680
 TTTACTAACG TCTGGAAAGA CGACAAAAGT TTAGATCGTT ACGCTAACTA TGAGGGTTGT 1740
 CTGTGGAATG CTACAGGGCGT TGAGTTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA 1800
 TGGGTTCTCA TTGGGTTGCA TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT 1860
 TCTGAGGGTG GCGGTTCTGA GGGTGGGGT ACTAAACCTC CTGAGTACGG TGATACACCT 1920
 ATTCCGGGCT ATACTTATAT CAACCCCTGTC GACGGCACTT ATCCGCTGG TACTGAGCAA 1980
 AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAAATAC TTTCATGTTT 2040
 CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACGT TTTATACGGG CACTGTTACT 2100
 CAAGGCAGTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG 2160
 TATGACCGCTT ACTGGAACGG TAAATTCAAGA GAGTGCCTT TCCATTCTGG CTTTAATGAA 2220
 GATCCATTG TTTGTGAATA TCAAGGCCAA TCGTGTGACC TGCCTCAACC TCCTGTCAAT 2280
 GCTGGCGGCG GCTCTGGTGG TGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT 2340
 GGCGGTTCTG AGGGTGGGGG CTCTGAGGGG GGCGGTTCTG GTGGTGGCTC TGTTCCGGT 2400
 GATTTTGATT ATGAAAAGAT GGCAAAACGCT ATAAGGGGG CTATGACCGA AAATGCCGAT 2460
 GAAAACGCCG TACAGTCTGA CGCTAAAGGC AAACCTGATT CTGTGCGCTAC TGATTACGGT 2520
 GCTGCTATCG ATGGTTTCAT TGTCACGTT TCCGGCCTTG CTAATGGTAA TGTTGCTACT 2580
 GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCAACCT 2640
 TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGTTGA ATGTCGCCCT 2700
 TTGTCCTTAA GCGCTGGTAA ACCATATGAA TTTCTATTG ATTGTGACAA AATAAAACTTA 2760
 TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTGCCACCT TTATGTATGT ATTTTCTAGG 2820
 TTGCTAACAA TACTGCGTAA TAAGGAGTCT TAATCAIGCC AGTTCTTTG GGTATTCCGT 2880
 TATTATTGCG TTTCTCGGT TTCCCTCTGG TAACCTTGTG CGGCTATCTG CTTACTTTG 2940
 TTAAAAAGGG CTTCGGTAAAG ATAGCTATTG CTATTCAATT GTTCTTGCT CTTATTATTG 3000
 GGCTTAACTC AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTAA CCCCTCTGACT 3060
 TTGTTCAAGGG TGTTCAAGTAA ATTCTCCGT CTAATGCCGT TCCCTGTTT TATGTATTC 3120
 TCTCTGTAAA GGCTGCTATT TTGATTGTTG ACGTTAAACA AAAAATCGTT TCTTATTG 3180
 ATTGGGATAA ATAATATGGC TGTTTATTT GTAACTGGCA AATTAGGCTC TGGAAAGACG 3240
 CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT 3300
 CTTGATTTAA GGCTTCAAAA CCTCCCGCAA GTCCGGAGGT TCGCTAAAAC GCCTCGCGT 3360
 CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGCG CGGTAATGAT 3420
 TCCTACCGATG AAAATAAAAAA CGGCTTGCTT GTTCTCGATG AGTGGGGTAC TTGGTTAAT 3480
 ACCCGTTCTT GGAATGATAA GGAAAGACAG CGCATTATTG ATTGGTTCT ACATGCTCGT 3540

AAATTAGGAT CGGATATTAT TTTTCTTGTGTT CAGGACTTAT CTATTGTTGA TAAACAGGGCG 3600
 CGTTCTGCAT TAGCTGAACA TGTGTTTAT TGTCTCGTC TGGACAGAAAT TACTTTACCT 3660
 TTTGTCGGTA CTTTATATTCTCTTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT 3720
 GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT 3780
 ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTCTAG TAATTATGAT 3840
 TCCGGTGTGTT ATTCTTATTT AACGCCCTAT TTATCACACCG GTCGGTATTT CAAACCATT 3900
 AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTC ACGGCTTCTT 3960
 TGTCTTGCAGTTCGATTCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG 4020
 GAGGTTAAAAA AGGTAGTCTC TCAGACCTAT GATTTGATA AATTCACTAT TGACTCTCT 4080
 CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAAT 4140
 AGCGACGATT TACAGAAGCA AGGTTATTCA CTACACATATA TTGATTTATG TACTGTTTCC 4200
 ATTAAAAAAAG CTAATTCAAA TCAAATTGTT AAATGTAATT AATTGTTT TCTTGATGTT 4260
 TGTTTCATCA TCTTCTTTG CTCAGGTAAATGAAAT AATTCCGCTC TGCGCGATTT 4320
 TGTAACCTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTTCTCCCG ATGTAAGG 4380
 TACTGTTACT CTATATTCACTGACCGTAAACCTGAAAT CTACCGCAATT TCTTTATTTC 4440
 TGTTTACGT GCTAATAATT TTGATATGGT TGCTTCAATT CCTTCATAA TTCAAGGTA 4500
 TAATCCAAAC AATCAGGATT ATATTGATGA ATTCGATCA TCTGATAATC AGGAATATGA 4560
 TGATAATTCC GCTCCTTCTG GTGGTTCTT TGTTCCGAA AATGATAATG TTACTCAAAC 4620
 TTTTAAAATT AATAACGTTG GGGCAAAGGA TTTAATACGA GTTCTCGAAT TGTTGTTAAA 4680
 GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT 4740
 TAGTGCACCT AAAGATATTG TAGATAACCT TCCTCAATC CTTTCTACTG TTGATTTGCC 4800
 AACTGACCAAG ATATTGATTG AGGGTTTGAT ATTTGAGGT CAGCAAGGTG ATGCTTITAGA 4860
 TTTTCACTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG 4920
 CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTGCTCGGT ATTTTAATG GCGATGTTT 4980
 AGGGCTATCA GTTCCGCGAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACCG 5040
 TATTCTTACG CTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCGAGAAATG TCCCTTTAT 5100
 TACTGGTCTG GTGACTGGTG AATCTGCCAA TGTAAATAAT CCATTCAAGA CGATTGAGGG 5160
 TCAAAATGTA GGTATTTCCA TGAGCGTTT TCCCTGTTGCA ATGGGTGGCG GTAATATTGT 5220
 TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTCT ACTCAGGCAA GTGATGTTAT 5280
 TACTAATCAA AGAAGTATTG CTACAACGGT TAATTGCGT GATGGACAGA CTCTTTACT 5340
 CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCCTGCTAA 5400
 AATCCCTTAA ATCGGGCTCC TGTGTTAGCTC CGGCTCTGAT TCCAACGAGG AAAGCAGGTT 5460
 ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGGCGATTAGCAGGAGG 5520
 GTGTGGTGGT TACGGCCAGC GTGACCGCTA CACTTGGCAG CGCCCTAGCC CCCCGCTCTT 5580

TCGCTTCTT CCCTCCCTT CTGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC 5640
 GGGGGCTCCC TTTAGGGTTC CGATTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG 5700
 ATTTGGTGA TGGTCACGT AGTGGCCAT CGCCCTGATA GACCGTTTT CGCCCTTGA 5760
 CGTTGGAGTC CACGTTCTT AATAGTGGAC TCTTGTCCA AACTGGAACA AACTCAACC 5820
 CTATCTCGGG CTATTCTTT GATTATAAG GGATTTGCC GATTCGGAA CCACCATCAA 5880
 ACAGGATTT CGCCTGCTGG GGCAAACCGAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG 5940
 CCAGGGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCCACCT 6000
 GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGCC GATTCACTAA TCCAGCTGGC 6060
 ACCACAGGTT TCCCGACTGG AAAGGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC 6120
 TCACTCATTAA GGCACCCCAAG GCTTACACT TTATGTTCC GGCTCGTATG TTGTGTGGAA 6180
 TTGTGAGCGG ATAACAATT TACACGCCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC 6240
 TACGGCAGCC GCTGGATTGT TATTACTCGG TCCCCAACCA GCCATGGCCG AGCTCTTCCC 6300
 GCCATCTGAT GAGCAGTTGA AATCTGGAAC TGCCTCTGTT GTGTGCCTGC TGAATAACTT 6360
 CTATCCCAGA GAGGCCAAAG TACAGTGGAA GGTGGATAAC GCCCTCCAAT CGGGTAACCTC 6420
 CCAGGAGAGT GTCACAGAGC AGGACAGCAA GGACAGCACC TACAGCCTCA GCAGCACCC 6480
 GACGCTGAGC AAAGCAGACT ACGAGAAACA CAAAGTCTAC GCCTGCGAAG TCACCCATCA 6540
 GGGCCTGAGC TCGCCCGTCA CAAAGAGCTT CAACAGGGGA GAGTGTCTA GAACGGTCA 6600
 CTTGCCACTG GCCGTCGTTT TACAACGTG TGACTGGAA AACCCCTGGG TTACCCAAGC 6660
 TTTGTACATG GAGAAAATAA AGTGAACCAA AGCACTATTG CACTGGCACT CTTACCGTTA 6720
 CTGTTTACCC CTGTGGCAAA AGCCCCCTCC ACCAAGGGCC CATCGGTCTT CCCCCCTGGCA 6780
 CCCTCCTCCA AGAGCACCTC TGGGGCACA GCGGGCCCTGG GCTGCCGTGGT CAAGACTAAT 6840
 TCCCCGAACC GGTGACGGTG TCGTCCAAT CAGGCGCCCT GACCAGGGCC GTGCACACCT 6900
 TCCCCGCTGT CCTACAGTCC TCAGGACTCT ACTCCCTCAG CAGCGTGGTG ACCGTGCCCT 6960
 CCAGCAGCTT GGGCACCCAG ACCTACATCT GCAACGTGAA TCACAAGCCC AGCAACACCA 7020
 AGGTGGACAA GAAAGCAGAG CCCAAATCTT GTACTAGTGG ATCCGTACCCG TACGACGTTC 7080
 CGGACTACGGC TTCTTACGGCT GAAGGGGATG ACCCTGCTAA GGCTGGCATTC AATAGTTTAC 7140
 AGGCAAGTGC TACTGAGTAC ATTGGCTACG CTTGGCTAT GGTAGTAGTT ATAGTTGGTG 7200
 CTACCATAGG GATTAAATTA TTCAAAAGT TTACGGAGCAA GGCTTCTTAA GCAATAGCGA 7260
 AGAGGCCCGC ACCGATCGCC CTTCCCAAGA GTTGGCGAGC CTGAATGGGG AATGGCGCTT 7320
 TGCCCTGGTTT CCGGCACCCAG AAGGGTGCC GGAAAGCTGG CTGGAGTGG ATCTTCTG 7380
 GGCCGATACG GTCGTCGTCC CCTCAAATG GCAGATGCAC GGTTACGGATG CGCCCATCTA 7440
 CACCAACGTA ACCTATCCCA TTACGGTCAA TCCGCCGTTC GTTCCCACGG AGAATCCGAC 7500
 GGGTTGTAC TCGCTCACAT TTAATGTTGA TGAAAGCTGG CTACAGGAAG GCGAGACGGG 7560
 AATTATTTT GATGGCGTTC CTATTGTTA AAAATGAGC TGATTTAAC AATTTAAC 7620

GGAAATTTA ACAAAATATT AACGTTTACA ATTTAAATAT TTGCTTATAC AATCTTCCTG	7680
TTTTGGGGC TTTCTGATT ATCAACCGGG GTACATATGA TTGACATGCT AGTTTACGA	7740
TTACCGTTCA TCGATTCTCT TGTTGCTCC AGACTCTAG GCAATGACCT GATAGCCTT	7800
GTAGATCTCT CAAAAATAGC TACCCCTCTCC GGCATTAATT TATCAGCTAG AACGGTTCAA	7860
TATCATATTG ATGGTGATT GACTGTCTCC GGCCTTCTC ACCCTTTGA ATCTTACCT	7920
ACACATTACT CAGGCATTGC ATTTAAATA TATGAGGGTT CTAAAAATT TTATCCTTGC	7980
GTTGAAATAA AGGCTTCTCC CGCAAAAGTA TTACAGGGTC ATAATGTTTT TGGTACAACC	8040
GATTTAGCTT TATGCTCTGA GGCTTTATTG CTTAATTTCG CTAATTCTTT GCCTTGCCCTG	8100
TATGATTAT TGGACGTT	8118

(2) INFORMATION FOR SEQ ID NO:6:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(5, "")
- (D) OTHER INFORMATION: /note- "S REPRESENTS EQUAL MIXTURE OF G AND C"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(6, "")
- (D) OTHER INFORMATION: /note- "M REPRESENTS EQUAL MIXTURE OF A AND C"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(8, "")
- (D) OTHER INFORMATION: /note- "R REPRESENTS EQUAL MIXTURE OF A AND G"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(11, "")
- (D) OTHER INFORMATION: /note- "K REPRESENTS EQUAL MIXTURE OF G AND T"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(20, "")
- (D) OTHER INFORMATION: /note- "W REPRESENTS EQUAL MIXTURE OF A AND T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGGTSMARCT KCTCGAGTCW GG

(2) INFORMATION FOR SEQ ID NO:7:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGGTCCAGCT GCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:8:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGTCCAGCT GCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:9:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGTCCAGCT TCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:10:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGGTCCAGCT TCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:11:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGGTCCAACT GCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGGTCCAACT GCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGGTCCAACT TCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGGTCCAACT TCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(5..6, "")
- (D) OTHER INFORMATION: /note- "N-INOSINE"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(8, "")
- (D) OTHER INFORMATION: /note- "N-INOSINE"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(11, "")
- (D) OTHER INFORMATION: /note- "N=INOSINE"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(20, "")
- (D) OTHER INFORMATION: /note- "W REPRESENTS EQUAL MIXTURE OF A AND T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AGGTNNANCT NCTCGAGTCW GG

22

(2) INFORMATION FOR SEQ ID NO:16:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTATTAACATA GTAACGGTAA CAGTGGTGCC TTGCCCCA

38

(2) INFORMATION FOR SEQ ID NO:17:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGCCTTACTA GTACAATCCC TGGGCACAAT

30

(2) INFORMATION FOR SEQ ID NO:18:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCAGTTCCGA GCTCGTTGTG ACTCAGGAAT CT

32

(2) INFORMATION FOR SEQ ID NO:19:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 32 bas pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: lin ar

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCAGTTCCGA GCTCGTGTG ACGCAGCCCC CC

32

(2) INFORMATION FOR SEQ ID NO:20:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCAGTTCCGA GCTCGTGCTC ACCCAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:21:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CCAGTTCCGA GCTCCAGATG ACCCAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:22:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCAGATGTCA GCTCGTGATG ACCCAGACTC CA

32

(2) INFORMATION FOR SEQ ID NO:23:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCAGATGTCA GCTCGTCATG ACCCAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:24:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: singl
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCAGTTCCGA GCTCGTGATC ACACAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCAGCATTCT AGAGTTTCAG CTCCAGCTTG CC

32

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCCCCGTCTA GAATTAACAC TCATTCCTGT TGAA

34

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GATCCTAGGC TGAAGGCGAT GACCCTGCTA AGGCTGC

37

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATTCAATAGT TTACAGCCAA GTGCTACTGA CTACA

35

(2) INFORMATION FOR SEQ ID NO:29:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTGGCTACGC TTGGGCTATG GTAGTAGTTA TAGTT

35

(2) INFORMATION FOR SEQ ID NO:30:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGTGCTACCA TAGGGATTAA ATTATTCAAA AAGTT

35

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TACCGAGCAAG GCTTCTTA

18

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AGCTTAAGAA GCCTTGCTCG TAAACTTTTT GAATAATT

39

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AATCCCTATG GTAGCACCAA CTATAACTAC TACCAT

36

(2) INFORMATION FOR SEQ ID NO:34:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

AGCCCAAGCG TAGCCAATGT ACTCACTAGC ACTTG

35

(2) INFORMATION FOR SEQ ID NO:35:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CCTGTAAACT ATTGAATGCA GCCTTAGCAG GGTC

34

(2) INFORMATION FOR SEQ ID NO:36:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36

ATCGCCTTCA GCCTAG

16

(2) INFORMATION FOR SEQ ID NO:37:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CATTTTTGGA GATGGCTTAG A

21

(2) INFORMATION FOR SEQ ID NO:38:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TAGCATTAAAC GTCCAATA

18

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ATATATTTTA GTAAGCTTCA TCTTCT

26

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GACAAAGAAC CGGTGAAAAC TTT

23

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GCGGGCCTCT TCGCTATTGC TTAAGAAGCC TTGCT

35

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AAACGACGGC CAGTGCCAAG TGACCCGTGT GAAATTGTTA TCC

43

(2) INFORMATION FOR SEQ ID NO:43:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGCGAAAGGG AATTCTGCAA GGCGATTAAAG CTTGGGTAAC GCC

43

(2) INFORMATION FOR SEQ ID NO:44:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGCGTTACCC AAGCTTTGTA CATGGAGAAA ATAAAG

36

(2) INFORMATION FOR SEQ ID NO:45:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TGAAACAAAG CACTATTGCA CTGGCACTCT TACCGTTACC GT

42

(2) INFORMATION FOR SEQ ID NO:46:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TACTGTTTAC CCCTGTGACA AAAGCCGGCC AGGTCCAGCT GC

42

(2) INFORMATION FOR SEQ ID NO:47:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TCGAGTCAGG CCTATTGTGC CCACGGATTG TACTAGTGGA TCCG

44

(2) INFORMATION FOR SEQ ID NO:48:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TGGCCAAAGG GAATTGGAT CCACTAGTAC AATCCCTG

38

(2) INFORMATION FOR SEQ ID NO:49:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GGCACAAATAG GCCTGACTCG AGCAGCTCGA CCAGGGCGGC TT

42

(2) INFORMATION FOR SEQ ID NO:50:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

TTGTCACAGG GGTAAACAGT AACGGTAACG GTAAAGTGTC CA

42

(2) INFORMATION FOR SEQ ID NO:51:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GTGCAATAGT GCTTGTTC ACTTATTT CTCCATGTAC AA

42

(2) INFORMATION FOR SEQ ID NO:52:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: singl
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:
TAACGGTAAG AGTGCCAGTG C

21

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:
CACCTTCATG AATTGGCAA GGAGACAGTC AT

32

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:
AATTGGCAA GGAGACAGTC AT

22

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:
AATGAAATAC CTATTGCCTA CGGGAGCCGC TGGATTGTT

39

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:
ATTACTCGCT GCCCAACCAG CCATGGCCGA GCTCGTGAT

39

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GACCCAGACT CCAGATATCC AACAGGAATG AGTGTAAAT

39

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCTAGAACCGC GTC

13

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTCAGGTTGA AGCTTACGGG TTCTAGAATT AACACTCAAT CCTGT

45

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TGGATATCTG GAGTCTGGGT CATCACGAGC TCGGCCATG

39

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: singl
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GCTGGTTGGG CAGCGAGTAA TAACAATCCA GCGGCTGCC

39

(2) INFORMATION FOR SEQ ID NO:62:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

GTAGGCAATA CGTATTTCAT TATGACTGTC CTTGGCG

37

(2) INFORMATION FOR SEQ ID NO:63:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TGACTGTCTC CTTGGCGTGT GAAATTGTAA

30

(2) INFORMATION FOR SEQ ID NO:64:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

TAACACTCAT TCCGGATGGA ATTCTGGAGT CTGGGT

36

(2) INFORMATION FOR SEQ ID NO:65:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GCCAGTGCCA AGTGACGCGT TCTA

24

(2) INFORMATION FOR SEQ ID NO:66:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATATATTTA GTAAGCTTCA TCTTCT

26

(2) INFORMATION FOR SEQ ID NO:67:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GACAAAGAAC GCGTGAAAAC TTT

23

(2) INFORMATION FOR SEQ ID NO:68:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CTGAACCTGT CTGGGACCAC AGTTGATGCT ATAGGATCAG ATCTAGAATT CATTAGAGA

60

CTGGCCTGGC TTCTGC

76

(2) INFORMATION FOR SEQ ID NO:69:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TCGACCGTTG GTAGGAATAA TCCAATTAAT GGAGTAGCTC TAAATTAGA ATTCACTAC

60

ACCCAGTGCA TCCAGTAGCT

80

(2) INFORMATION FOR SEQ ID NO:70:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

GGTAAACAGT AACGGTAAGA GTGCCAG

27

(2) INFORMATION FOR SEQ ID NO:71:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

CGCCTTCAGC CTAAGAACGG TAGTCCGGAA CGTCGTACGG GTAGGATCCA CTAG

54

(2) INFORMATION FOR SEQ ID NO:72:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CACCGGTTCG GGGATTAGT CTTGACCAGG CAGCCCAGGG C

41

(2) INFORMATION FOR SEQ ID NO:73:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

ATCCACACA TTATACCGAGC CGGAAGCATA AAGTGTCAAG CCTGGGGTGC C

51

(2) INFORMATION FOR SEQ ID NO:74:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CTGGTCATCA GATGGCGGGA AGAGCTCGGC CATGGCTGGT TG

42

(2) INFORMATION FOR SEQ ID NO:75:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: singl
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:
GAACAGAGTG ACCGAGGGGG CGAGCTGGGC CATGGCTGGT TG